

APPLICATION
FOR
UNITED STATES LETTERS PATENT

APPLICANT NAME: S. E. Aschen et al.

TITLE: FILE RETRIEVAL METHOD AND SYSTEM

DOCKET NO.: END920030038US1

INTERNATIONAL BUSINESS MACHINES CORPORATION

Certificate of Mailing Under 37 CFR 1.10

I hereby certify that, on the date shown below, this correspondence is being deposited with the United States Postal Service in an envelope addressed to the Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 as "Express Mail Post Office to Addressee"

"Express Mail" Label No.: EV 342658763 US

On: 6/25/03

Denise M. Jurik

Typed or Printed Name of Person Mailing Correspondence

Denise M Jurik 06/25/03
Signature Date

FILE RETRIEVAL METHOD AND SYSTEM

CROSS REFERENCE TO RELATED APPLICATION

This application is related to an application titled "MAIL
AND CALENDAR TOOL AND METHOD" having common inventors and a
5 common assignee. Both applications are filed on the same date.

TECHNICAL FIELD

The invention relates to providing data stored in a database
by a first application to other applications. More particularly
the invention relates to providing such data from a database
10 residing on a server to applications residing on other servers or
client computers. Even more particularly, such data is provided
to the other applications without the need or use of the first
application, and in a simple object access protocol.

BACKGROUND OF THE INVENTION

15 Various systems and applications have been developed to
provide electronic mail (e-mail) and calendar maintenance
functions to users. Some systems make use of an application
running on a user workstation, such as OUTLOOK EXPRESS®
available from Microsoft Corp. (OUTLOOK EXPRESS is a trademark
20 of Microsoft Corp. of Redmond, WA). The application retrieves
incoming e-mail through an internet connection and displays a
list of incoming notes. The user may select one or more notes to
read, delete, save, or forward to another user. Similarly the

application allows a user to compose an outgoing e-mail note and send it through the same internet connection. copies of the incoming and outgoing notes may generally be saved on either the user's workstation, or on a server supporting the internet connection, or both.

Other systems use an ordinary internet browser running on a user workstation to provide similar functions. Still others may use an ordinary browser with some modification to facilitate e-mail and calendar maintenance (calendar) functions.

Some applications are specifically developed to provide e-mail and calendar for numerous employees in a company. LOTUS NOTES® available from International Business Machines Corp. (LOTUS NOTES is a trademark of International Business Machines Corp.) is one example of such an application. The individual notes and calendar entries are stored in a database running on a server such as the DOMINO® server software (DOMINO is a trademark of International Business Machines Corp.) a copy or replica of the stored notes and calendar entries may also be kept on a user's workstation to permit standalone operation, for example when the connection to the server, or the server itself is unavailable due to overload or breakdown.

Some of the information contained in these notes and calendar entries may be extremely valuable, particularly in the case of a large company where critical business information may be held by numerous employees at various locations. Other business applications may benefit from having access to the information kept in this server database, however no means is readily available to provide such access without extensive coding effort by the developers of the other application. For example,

a note or calendar entry document may have sections, attachments, image tags, and links to other items in the document. Ordinary document retrieval using a browser will not make these sections, attachments, images, or links usable to the application using a browser for retrieval.

While data stored by a mail and calendaring application is of primary interest, data stored by any application may be of value to other applications if the data can be made visible in a readily discernable manner.

It would therefore be a significant accomplishment if a system or method were developed to easily provide such information to other applications. Furthermore, other applications may be running on computer systems which do not have the first application installed or available for use. It would therefore be a desirable feature to provide such information without making use of the first application. It is believed that this would constitute a significant advancement in the data retrieval arts.

OBJECTS AND SUMMARY OF THE INVENTION

It is therefore a principal object of the present invention to enhance the data retrieval art by providing a method of access with enhanced capability.

It is another object to provide a system with such enhanced capability.

These and other objects are attained in accordance with one embodiment of the invention wherein there is provided a method of exposing a file to an application, comprising the steps of, providing a file of documents having fields, receiving a request
5 for one or more of the fields of one of the documents as a message in a simple object access protocol from an application, extracting the one or more of the fields from the file as an extended markup document, parsing the extended markup document according to a schema, authenticating the application, and
10 sending the parsed document as a simple object access protocol message to the application.

In accordance with another embodiment of the invention there is provided a method of providing data to an application, comprising the steps of, providing a mailfile of documents having
15 a section and fields, receiving a request as a SOAP protocol message from an application for one of the documents, retrieving the fields of the one of the documents from the mailfile, in response to the fields, retrieving the one of the documents as a markup language document, inserting a URL into the markup
20 language document to retrieve the section of the one of the documents, retrieving the section from the mailfile in the markup language, removing the URL from the retrieved document and creating an object having the section expanded in the retrieved document, and marshalling the object and sending the marshalled
25 object to the application as a SOAP protocol message.

In accordance with yet another embodiment of the invention there is provided a system for providing data stored in a file to an application, comprising, a file having data stored as documents, a database for passing a request for one of the
30 documents to the file and upon return converting the one of the

documents into an extended markup format, an authentication directory having authentication records for an application, web service software for receiving a request from an application for one of the documents, retrieving the one of the documents, and
5 creating an extended markup object containing the document, and a protocol tool for authenticating the application using the records, marshalling the object, and sending the marshalled object in a simple object access protocol to the application.

BRIEF DESCRIPTION OF THE DRAWINGS

10 FIG. 1 depicts a mail and calendaring system; and

FIG. 2 is a document retrieval flowchart.

BEST MODE FOR CARRYING OUT THE INVENTION

For a better understanding of the present invention, together with other and further objects, advantages and capabilities thereof, reference is made to the following
15 disclosure and the appended claims in connection with the above-described drawings.

In FIG. 1 there is shown elements of a system for providing mail and calendaring data stored in mailfile 12 to application 18
20 in accordance with the present invention. Mailfile 12 contains documents stored by a mail and calendaring client (not shown). For example, a client mail and calendaring program product such

as LOTUS NOTES may store each e-mail note or calendar entry for a user as a document in mailfile 12. Understandably mailfile 12 may be very large containing thousands or millions of documents, particularly if many employees in a large company share the mailfile.

As noted above any application may save data in a file, in which case mailfile 12 may contain data from any such application.

The documents have fields and at least one section. A section is defined as additional data which may be hidden or visible when the document is viewed. For example, the document may have a subtitle, category, or other term with an expansion button nearby. A triangle shaped "twistee" button or any other type of expansion control may be used. When the expansion control is activated, such as by clicking a mouse pointer on a twistee, the additional data is either exposed or hidden from view.

The document may also include an attachment link to an attachment stored in mailfile 12, or stored elsewhere on the same server where mailfile 12 is located, or stored elsewhere in a network.

The document may also include an image tag. The image itself may also be stored elsewhere. Any format of image data whether coded or uncoded may be used such as joint photographic experts group (JPEG or .JPG), bitmap (BMP), graphics interchange (.GIF) or any other image format.

The document may also include a link to other objects. The link may point to information stored elsewhere.

Application 18 requests access to a document stored in mailfile 12. The request is sent to mail and calendaring web service software 16 over communication link 19. Any type of communication protocol may be used for sending and receiving the request including hypertext transfer protocol (HTTP), or HTTP secure (HTTPS) either alone or as a carrier for a simple object access protocol (SOAP) message. Application 18 may include a client 20 for handling a SOAP message.

Authentication records for application 18 are stored in authentication directory 22 which may be any type of directory such as a DOMINO directory or an LDAP (Lightweight Directory Access Protocol) directory or any other type of directory known in the art. Directory 22 may also be a password file or a credential vault.

Database 14 is shown in FIG. 1 as residing on the same server as mailfile 12, however this may be a replica mailfile as shown. For example, mailfile 12 and database 14 may reside on DOMINO server. Database 14 has a capability of requesting a list of available documents from mailfile 12 or of passing a request for a document to mailfile 12. When the document is returned, database 14 converts the documents into an extended markup format such as XML.

Mail and calendaring web service software (MACS) 16 receives the document request from application 18 over communication link 19. MACS retrieves the document from mailfile 12 in accordance with the retrieval flowchart of FIG. 2.

In step 40 of FIG. 2 the document fields are retrieved from mailfile 12 through use of database 14. Fields of the document

END920030038US1

are retrieved in XML format. However, as noted above, the document may have unexpanded sections which are not visible to the requesting application. Consequently loop steps 44, 46, 48, and 50 access mail database 12 to obtain the sections and insert these into the retrieved document. A URL for each section is created and inserted in the retrieval document so the section can be obtained, for example with an HTML Get request from mailfile 12. The sections are inserted in the retrieval document at the appropriate point. Later, in step 58 these URL's are removed when no longer needed.

Non-body HTML statements such as header, are removed in step 52.

In steps 54 and 56, attachments are obtained and inserted by MACS 16 in place of attachment links in the document. In step 60 images are obtained and inserted in the document in response to image tags. The image data so obtained may need to be encoded in a format compatible with the markup language e.g. XML, HTML used for the text portion of the document and any restriction imposed by the transmission protocol. One such compatible encoding, known as Base 64 as defined in section 6.8 (page 23) of the MIME part 1 document may be used. Other encodings may also be used.

In step 62 iframe links in the retrieved document are removed by MACS 16 and the iframe content is inserted into the document.

In step 64 an XML object of the entire document is created by MACS 16.

Returning now to FIG. 1, mailfile 12 and database 14 have access to authentication directory 22 as shown. Whenever data is

END920030038US1

retrieved from mailfile 12 for the purpose of satisfying a request from application 18, an authentication is performed using the records stored in directory 22.

5 XML object of step 64 is then returned to requesting application 18 over communication link 19. Any communication protocol may be used including HTTP alone, secure, or as a carrier to a SOAP message.

10 The object is converted or marshalled into a SOAP message according to a pre-defined set of rules. Because the object has a complex structure comprising XML text, inserted sections, attachments, images, and iframes, a complex set of rules is needed. The rules are formulated to operate in conjunction with MACS 16, which creates the complex object.

15 The SOAP message may then be sent over communication link 19 to requesting application 18 using any transmission protocol. In a preferred embodiment the HTTPS protocol is used. Application 18 may then extract the data from the SOAP message using SOAP client 20, and use the data in any manner required.

20 While there have been shown and described what are at present considered the preferred embodiments of the invention, it will be obvious to those skilled in the art that various changes and modifications may be made therein without departing from the scope of the invention as defined by the appended claims.